What Causes Heart Disease, Kidney Disease, and Skin Cancer?

Kate M. Sebaganji, Ndeh M. Awasum, Carynn Vuong

*STAT 360*

ABSTRACT

**Introduction:**

Heart disease, skin cancer, and kidney disease are some of the most prevalent diseases and top leading causes of death in the US. CDC reports that one person dies every 34 seconds from cardiovascular disease, and one in 5 five Americans will develop skin cancer in their lifetime. Given the fact that these diseases are very likely, we used structural model equation (SEM) to analyze the biggest causes of such diseases, quantify their effects, and understand the comprehensive relationships among the predictor variables. This model will contribute towards a better understanding of how people can maintain healthy lifestyles to avoid cancer and heart diseases.

**Methods:**

The dataset used in this research was collected by the CDC in their surveys about the health status of Americans across all US territories. After cleaning the data, assigning numerical values to categorical variables, and removing any cases with NA values, our dataset had 319,795 cases, 14 predictor variables and 3 response variables. Based on their relevance to our study, the retained predictor variables are smoking, BMI, alcohol\_drinking, sex, asthma, sleep\_time, diabetic, mental\_health, stroke, general\_health, physical\_activity, physical\_health, difficulty\_walking, and average\_age. We used Kaiser’s and Jolliffe’s criterion to establish a loading matrix that was used to create a path diagram from which we derived the SEM for our data. We later evaluated the fitness of our model to justify its relevance and applicability to our original data.

**Results:**

Apart from the non-significant influence of mental health on the prevalence of kidney disease, our model proved that all the evaluated predictor variables have a statistically significant influence on the 3 diseases. Stroke had the biggest effect on both heart and kidney diseases; as the stroke occurrence increases by 1 unit, the chance of getting a heart or kidney disease increases by 0.192 units and 0.044 units respectively. Our analysis also shows that being female instead of male leads to a 0.021 unit decrease in the chance of having skin cancer.

**Conclusion:**

In terms of the fitness of our model, our model has a reasonable fit according to the GFI and RMSEA indices. Approximately, 94.8% of the variance and covariance in the original data matrix can be explained by the structural equation model. Our model provides a good starting point to analyze the most impactful causes of heart disease, skin, and kidney cancers. Further research could focus on analyzing how socio-economic factors contribute to the studied diseases’ prevalence.

